Письма в ЖЭТФ, том 56, вып.9, стр.497 - 501 10 ноября 1992 г.

РАСПОЗНАВАНИЕ ОБРАЗЦОВ ЧЕРЕЗ МОДЕЛЬ ДЕРРИДЫ С РЕДКИМИ СВЯЗЯМИ

Д.Б.Саакян

Ереванский физический институт, Армения

Поступила в редакцию 22 апреля 1992 г.
После переработки 2 октября 1992 г.

Рассматривается модель типа Дерриды с редкими связями для распознавания образов. Найдены критические значения связанности и числа образов, при котором происходит распознавание образов.

Модель Дерриды дает простейший образец состояния спинового стекла. Точное решение для нее получается уже при однократном нарушении репличной симметрии в рамках схемы Паризи [2]. Поэтому, пожалуй, система обладает рядом уникальных свойств, например, дает возможность оптимальным образом кодировать информацию при наличии помех [3].

В [4] использовалось для распознавания образов полностью связанный вариант модели.

В настоящей статье продолжим идеологию [4] для случая редких связей. При этом будем использовать технику [5] для моделей с редкими связями. Пусть задано M случайных образов

$$\{\xi_c^\mu\}, \quad (\mu = 1...M, \ i = 1...N), \qquad \xi_i^\mu = \pm 1.$$

Мы хотим построить такой гамильтониан H для N спинов σ_i, единственным основным состоянием которого являлись бы M наборов спинов $\{\xi_i^\mu\}$. Рассмотрим

$$H = -\sum_{(i_1..i_p)} \sigma_{i_1}..\sigma_{i_p} j_{i_1..i_p}, \tag{1}$$

где

$$j_{i_1..i_p} = \sum_{\mu=1}^{M} \xi_{i_1}^\mu..\xi_{i_p}^\mu. \tag{2}$$

В (1) идет суммирование не по всем возможным C_N^p выборам совокупности разных индексов $(i_1..i_p)$, а лишь $Z = cN$ из них, выбранных случайным образом. В рассмотренном случае мы берем $1 \ll p \ll N$ и $C_N^p \approx N^p/p!$. В модели Дерриды берется гамильтониан (1), только константы взаимодействия $j_{i_1..i_p}$ распределены по Гауссу [1] или дискретным образом [5]. Мы можем считать, что в (1) идет суммирование по всем наборам индексов $(i_1..i_p)$ без ограничения, только теперь константы $j_{i_1..i_p}$ имеют следующее распределение

$$\rho(j_{i_1..i_p}) = (1 - \frac{cp!}{N^{p-1}})\delta_{j_{i_1..i_p},0} + \frac{cp!}{N^{p-1}}\delta_{j_{i_1..i_p}},(\sum_{\mu=1}^{M} \xi_{i_1}^\mu..\xi_{i_p}^\mu). \tag{3}$$

Последнее представление мы используем, когда будем вычислять методом среднего поля свободную энергию модели.

Отметим, что ввиду калибровочной инвариантности системы можно вместо образов $\{\xi_i^\mu\}$ рассмотреть новые образы (1) и $\{\xi_i^1\xi_i^\mu\}$, $\mu \geq 2$ для которых (2) переходит в

$$j_{i_1..i_p} = 1 + \sum_{\mu=2}^{\mu}(\xi_{i_1}^1\xi_{i_1}^\mu)..(\xi_{i_p}^1\xi_{i_p}^\mu). \tag{4}$$

При рассмотрении ферромагнитной фазы мы будем использовать представление (4). Будем использовать формулу:

$$e^{jB\sum_{\alpha=1}^{n}\tau\sigma_\alpha} = (\mathrm{ch}Bj)^n[1 + \sum_{r=1}^{\infty}(\mathrm{th}Bj)^r \sum_{(\alpha_1..\alpha_r)} \sigma_{\alpha_1}..\sigma_{\alpha_r}\tau^r], \tag{5}$$

где в правой части идет суммирование по всем возможным наборам $(\alpha_1..\alpha_r)$ разных индексов. Рассматривая n копий спинов с гамильтонианом (1) можно получить (смотри вывод в [5]) для свободной энергии

$$-nBF = cn < \ln(\mathrm{ch}\, Bj) > - \sum_{r=1}^{\infty}\sum_{\alpha_1..\alpha_2} \lambda_{\alpha_1..\alpha_r}Q_{\alpha_1..\alpha_r} +$$

$$+c\sum_{r=1}^{\infty} < [\mathrm{th}Bj]^r > \sum_{\alpha_1..\alpha_r}(Q_{\alpha_1..\alpha_r})^p + \ln \mathrm{Tr}\exp\left[\sum_{r=1}^{\infty}\lambda_{\alpha_1..\alpha_r}\sigma_{\alpha_1}..\sigma_{\alpha_r}\right], \tag{6}$$

где $\lambda_{\alpha_1..\alpha_2}$ являются лагранжевыми множителями, а $Q_{\alpha_1..\alpha_r}$ – корреляторами спинов. Случайная величина определена как сумма M случайных чисел $\xi_i^\mu = \pm 1$

$$j = \sum_{\mu=1}^{M}\xi^\mu, \tag{7}$$

где ξ^μ имеет тот же закон распределения, что и ξ_i^μ для фиксированного i.

Далее в настоящей статье будем рассматривать случай, когда величины ξ_i^μ распределены равномерно. Полученные результаты легко обобщаются для неравномерного распределения ξ_i^μ. При рассмотрении ферромагнитной фазы как в (4), будем исходить из

$$j = 1 + \sum_{\mu=2}^{M}\xi^\mu, \tag{8}$$

где ξ^μ распределены равномерно.

Рассмотрим сначала решения (6) с сохранением репличной симметрии. Имеем

$$\lambda_{\alpha_1..\alpha_r} = \lambda_r, \qquad Q_{\alpha_1..\alpha_r} = q_r.$$

Из условия экстремума свободной энергии по q_r получаем из (6)

$$\lambda_r = cp < (\mathrm{th}Bj)^r > q_r^{p-1}. \tag{9}$$

Поскольку $q_r \leq 1$, то возможны ситуации

$$q_r = 1, \qquad \lambda_r \to \infty \tag{10}$$

и

$$q_r < 1, \qquad \lambda_r \to 0. \tag{11}$$

Парамагнитная фаза соответствует выбору (11). Разлагая $\ln \mathrm{Tr}_\sigma \exp \sum\limits_r \lambda_2 \sum\limits_{\alpha_1..\alpha_r} \sigma_{\alpha_1}..\sigma_{\alpha_r}$ в ряд получим, что ненулевым являются слагаемые с $(\lambda_r)^2$ и выше. Поэтому получим дифференцируя (6) по λ_r, что $q_r = 0$. Отсюда получим для свободной энергии парамагнитной фазы

$$-BF = c < \ln(\mathrm{ch} Bj) > + \ln 2. \tag{12}$$

Перейдем к ферромагнитной фазе. Пусть отлична от нуля q_1 (обычная намагниченность). Если бы было $q_r < 1$, то из (11) мы вернулись бы опять к ситуации парамагнетизма. Если же положить $q_r = 1$, то мы получим, из (10), что $\lambda_r \to \infty$. Тогда с точностью до членов малых как $\exp(-\lambda_1)$

$$\ln \mathrm{Tr}_\sigma \exp(\sum_r \lambda_r \sum_{\alpha_1..\alpha_r} \sigma_{\alpha_1}..\sigma_{\alpha_r}) = \sum_r \lambda_r C_n^r). \tag{13}$$

Используя (13) в (6) получаем из условия экстремума по λ_r, что для всех

$$q_r = 1 \tag{14}$$

и соответственно

$$\lambda_r \to \infty. \tag{15}$$

Подставляя (14), (15) в (6) получим при $n \to 0$

$$-nBF = cn \sum_{r=1}^{\infty} < \ln(\mathrm{ch} Bj) > + c \sum_{r=1}^{\infty} < (\mathrm{th} Bj)^r > C_n^r =$$

$$= cn < \ln(\mathrm{ch} Bj) > + c < [1 + \mathrm{th}(Bj)]^n - 1 > = cnB. \tag{16}$$

Мы использовали представление (8). В принципе могли бы быть (чисто гипотетические) решения, в которых $q_1 = 1$, $q_r = 1$ для некоторых r, но мы не видим никаких физических оснований для этого, в фазе сохранения репличной симметрии. Перейдем к фазе спинового стекла. При этом, оказывается, что для нечетных r $q_r = 0$ (а то мы опять оказались бы в фазе ферромагнетизма). Поэтому, мы имеем дело абсолютно с той же ситуацией, что и в [5], где рассматривается модель со случайным симметричным распределением констант $j_{i_1 - i_p}$. Как показано в [5], точное решение фазы спинового стекла получается уже при однократном нарушении репличной симметрии (рассматривание высших порядков нарушения репличной симметрии оставляет найденное решение без изменений), и значение для свободной энергии можно получить из соответствующего решения для парамагнитной фазы, рассматривая некую критическую B_c, при котором исчезает энтропия.

Отсюда имеем уравнение

$$c < \ln(\mathrm{ch} B_c j) > + \ln 2 - c < j \mathrm{th}(B_c j) > B_c = 0. \tag{17}$$

Начиная с этой температуры и выше свободная энергия не меняется. Поэтому для него (в спинстекольной фазе) имеем

$$-B_c F \equiv c < \ln(\mathrm{ch} B_c j) > + \ln 2 = c B_c < j \mathrm{th}(B_c j) > . \tag{18}$$

Мы нашли три фазы для нашей системы. В отличии от случая конечного p отсутствуют фазы с частичной намагниченностью. Такая ситуация с фазами в случае полной связанности [4] и не меняется в нашем случае редких связок. При каких значениях параметров (c, M, B) происходит переход между этими фазами? Математически строго надо было бы исследовать устойчивость найденных решений относительно возмущений. Но в нашей ситуации с случайными связями это очень сложно (если вообще возможно). С другой стороны ясно, что все эти три фазы должны существовать, во всяком случае при удачных значениях параметров.

Ферромагнитная фаза существует при $c \gg 1$, $B \gg 1$, маленькой парамагнитная при $B \ll 1$, спин стекольная при $B \gg 1$, $c \gg 1$, $M \gg 1$. Единственно, что нам остается, это найти границы фаз, и это мы сделаем из условия совпадения свободных энергий.

Когда происходит переход из фазы ферромагнетизма в фазу спинового стекла? Из условия сшивки решений (16), (18) получаем

$$1 = < j\,\mathrm{th}(B_c j) > . \tag{19}$$

Для $M = 2$ имеем из (19)

$$1 = \mathrm{th}(2B_c). \tag{20}$$

Поэтому $B_c \to \infty$ и получаем для критического c из (17)

$$c = 2. \tag{21}$$

Таким образом, при запоминании двух образов требуется $2N$ связок. В случае $M > 2$ требуется уже $Z > MN$ связок.

Если представим распределение j из (8) через

$$\rho(j) = \sum_{k=-M}^{M} \rho_k \delta_{k,j}, \tag{22}$$

то уравнения (19), (17) примут вид

$$1 = \sum_{k=-M}^{M} \rho_k \mathrm{th}(kB_c) \tag{23}$$

$$c = \frac{\ln 2}{\sum\limits_{k=-M}^{M} \rho_k [kB_c \mathrm{th}(B_c k) - \ln(\mathrm{ch} kB_c)]}. \tag{24}$$

Находя B_c из (23) и подставляя в (24), найдем критическое c. В пределе когда $M \to \infty$, система (23), (24) упрощается, и мы находим:

$$B_c = 1/M, \qquad c = M \ln 2. \tag{25}$$

Мы видим, что наш гамильтониан (1) запоминает образы не оптимальным способом. Очень интересно было бы найти такой гамильтониан, который, как и в случае кодирования, запоминал бы M образов оптимальным способом.

Рассмотрим теперь (1) в случае полной связанности. Существует любопытный эффект, фактически найденный в [4].

В [6] было введено понятие траты свободной энергии для выбора вакуума во время фазовых переходов в статистической физике. Если $Z(N,T)$ является статсуммой модели на решетке с N узлами при температуре T, и при некой критической температуре происходит фазовый переход, то терпит скачок асимптотика свободной энергии. Если существует асимптотическое разложение

$$F(N,T) \equiv \ln Z(N,T) = N f_0(T) + \ldots f_1(T) + \ldots O(1), \tag{26}$$

то

$$\lim_{\epsilon\to 0} f_1(T_c-\epsilon) - f_1(T_c+\epsilon) = \mathrm{const} \ln Q. \tag{27}$$

Су ествование скачка в (27) было известно давно из точного решения модели Изинга на конечной квадратичной решетке, где $\mathrm{const} = 1$, а порядок группы симметрии $Q = 2$.

В случае решеток с топологией тора, как показали численные расчеты [6], в (27) $\mathrm{const} = 1$ для $d2$, $Q = 3$, $Q = 5$ моделей Поттса и для $d3$-модели Изинга, как и для модели с сингулярностью Ли-Янга на случайной динамической решетке. Для $d2$-модели Изинга на случайной динамической решетке с родом g $\mathrm{const} = g$. В нашем случае для M образцов получаем, что

$$\lim_{\epsilon\to 0} f_1(T_c-\epsilon) - f_1(T_c+\epsilon) = \ln M. \tag{28}$$

Для того, чтобы выбрать нужный образ среди M возможных, система увеличивает свою энтропию на $\ln M$. В нашем случае гамильтониан не обладает точной группой симметрии (энергия разных вакуумов отличаются на $O(1)$ (в отличие от моделей из [6–8]. По-видимому существует свойство универсальности (27).

1. B.Derrida, Phys. Rev. Lett. **45**, 79 (1981).
2. D.Gross and M.Mezard, Nucl. Phys. **B240**, 141 (1984).
3. Д.Б.Саакян, Письма в ЖЭТФ, **55**, 198 (1992).
4. E.Gardner, J. Phys. A **20**, 3453 (1987).
5. C.Dominic and P.Mottishaw, J. de Phys. A **20**, L1267 (1987).
6. Д.Б.Саакян, Письма в ЖЭТФ, **52**, 717 (1990).
7. D.Sahakian, Modern Phys. Lett. A, n.2 (1992).
8. D.Sahakian, Phys. Lett. B **263**, n.3 (1991).